



## NATURAL AND INTUITIVE GESTURE INTERACTION FOR 3D OBJECT MANIPULATION IN CONCEPTUAL DESIGN

Tijana Vuletic, Alex Duffy, Laura Hay, Chris McTeague, Gerard Campbell, Pei Ling Choo,  
Madeleine Grealy

### Abstract

Gesture interaction with three-dimensional (3D) representations is increasingly explored, however there is little research present on the nature of the gestures used. A study was conducted in order to explore gestures designers perform naturally and intuitively while interacting with 3D objects during conceptual design. The findings demonstrate that different designers perform similar gestures for the same activities, and that their interaction with a 3D representation on a 2D screen is consistent with that which would be expected if a physical object were suspended in air in front of them.

*natural interaction, gestures, 3D modelling, conceptual design, case study*

### 1. Introduction

In an attempt to achieve more natural and intuitive ways of interaction with three-dimensional (3D) objects ubiquitous in computer-supported visualisation, interaction interface modalities alternative to WIMP (Window Icon Mouse Pointer) are widely explored. Hand gestures are particularly popular due to their naturalness and inherent presence in human communication processes (McNeill, 1985, Zimmerman et al., 1987, Buchmann et al., 2004, Quek, 2004). Gesture interface research is extensive, and includes use of gestures for virtual sculpting (Gao and Gibson, 2006, Han and Han, 2014, Vinayak and Ramani, 2015), interaction with large displays (Asadzadeh et al., 2012), architectural urban planning (Buchmann et al., 2004) and manipulation of objects in augmented reality (AR) (Hürst and van Wezel, 2013), virtual reality (VR) (Kim and Park, 2014) or 3D environments (Yeo et al., 2015). However, the focus is usually interface implementation investigation, and not the nature and appropriateness of the gestures for a specific use. Some of the technologies used are gloves (Dani and Gadh, 1997), cameras (Hürst and van Wezel, 2013), depth sensors (Song et al., 2014, Beattie et al., 2015), and motion capture systems (Han and Han, 2014). Occasionally these are supported by secondary technologies such as haptic systems (Zhu, 2008), gaze tracking (Song et al., 2014), electroencephalography/electromyography (EEG/EMG) (Xu et al., 2009, Shankar and Rai, 2014), or speech recognition (Bourdote et al., 2010). The gestures used for interaction are often highly influenced by the capabilities of the technologies facilitating their use (Schmidt, 2015). In some instances, gestures are initially generated by the users (Wright et al., 2011), and then a number of gestures are chosen to be adopted for the activities which are to be performed. While user generated gestures introduce an element of inherent user evaluation of gestures, in these cases, the final gestures are most frequently chosen based on the number of repetitions of a particular gesture, rather than the nature of the gesture itself and its appropriateness for the observed activity.

In the field of design research, the use of gestures in design is being explored independently of interface development. Nevertheless, even design research field rarely focuses on in-air, 3D hand gesture elicitation, its methodology and frameworks. Some authors focus on the role of gestures in a group collaboration. Cash and Maier (2016) have observed natural gestures designers use when collaborating within a group during a process of establishment of the shared understanding, and found gestures play a significant role. Others, like Hurtienne et al. (2010), have taken a different approach and found that primary metaphors can be applied to 3D freeform gestures, observed in terms of interaction design. However, majority of gesture systems for model creation, manipulation and modification either use prescribed gestures (Arroyave-Tobón et al., 2015, Noor and Aras, 2015), or copy free-form paths or surfaces traced by a hand (Qin et al., 2006, Vinayak and Ramani, 2015), that have not been extensively evaluated in terms of appropriateness for use in specific applications. Further research is needed to establish the most appropriate, intuitive and natural hand gestures for design activity. This has been explored for table top (2D surface) gestures where large sets of participants were involved in gesture generation, and it was found that the users generate more diverse gestures than those gesture designers have defined (Wobbrock et al., 2009). It was also found that users prefer user designed gestures that were, unbeknownst to them, suggested by a larger number of people, and gravitate towards physically and conceptually simpler gestures (Morris et al., 2010). While these findings are interesting, they cannot be transposed to in-air/3D free-form hand gestures for conceptual design, and similar research in this field has not been performed.

This paper reports on a gesture study investigating the natural and intuitive hand gestures designers perform when interacting with a 3D object, manipulating it or modifying it. The motivation behind this research is to explore the methodology for gesture elicitation which would help identify the most appropriate gestures for support of conceptual design, enabling more natural use of interfaces. The hand gestures used are dynamic, free-form, hand gestures, performing unrestricted in-air/3D motion. Study participants were free to create their own gestures for a number of design activities, primarily manipulative and a small number of modification based activities, assumed to be performed during the conceptual design stage. They were not instructed to perform a specific type of gestures, but they were asked to perform the gestures in order to achieve specific predefined activities. The activities chosen for inclusion in the study focus on conceptual design stage, where detail is not fully defined, and manipulation and modification of the objects is frequent. The goal was to identify the intuitive and instinctive response to stimuli. They first observe and then attempt to perform a number of activities using gestures (finger, hand or arm). The designers were working alone, and the focus was on user generated 3D free-form gestures. Technology used to facilitate the study consisted of a 2D screen, a camera and a LEAP sensor. A screen and a camera are technologies present in most workplaces and if the results show that they are appropriate for gesture identification that would mean similar studies designed to identify natural and intuitive gestures for 3D object interaction could be ran without extensive investment in the portable equipment. This is a preliminary exploratory study aiming to test the validity of the approach and provide initial findings on instances of repetition of same gestures, both between and within the participants.

Findings stemming from testing of the hypothesis listed in the following section should provide preliminary answers to the research questions:

- Can a 2D screen be used for the visualisation of the 3D activities without influencing the findings significantly?
- Is there repetition between gestures participants perform to complete activities without specific instructions given to them?

## 2. Hypotheses

Four hypotheses are stated to be tested by the study:

1. There is repetition between subjects for the same activities.  
Study results will be analysed to identify if different participants perform same gestures for the same activity, independently and without guidance or limitations. For example, if different participants use the same gesture to zoom into an observed object.
2. There is repetition within subjects for the same type of gestures.

Study results will be analysed to identify if a designer uses the same gesture to perform the same type of activity. For example, if they use the same gesture to rotate two different 3D objects at two different time instances.

3. 3D representation shown on a 2D screen does not significantly affect the perception of the 3D object.

Study results will be analysed to determine if participants interact with a 3D object shown on a 2D screen as if it was a physical object suspended in the air in front of them.

4. Type of the 3D object being manipulated (its recognisability) affects the gesture.

Study results will be analysed to determine if designers grasp 3D models of different objects in a different manner. For example, do they grasp a 3D model of a phone when they rotate it the same way they grasp a spherical 3D object.

### 3. Study methodology

An animation of a 3D object being manipulated or modified was shown to the participants three times in a row on a large screen. They were asked to observe it the first two times, and then to concurrently perform gestures using their hands the third time they saw it. The time between the second observation of the activity and the gesture performed was three seconds, and each animation lasted three seconds. This was done in order to record the initial reaction participants have, without allowing extra time which could be used to think about the activity they are performing or create an analogy with the way they would perform the same activity using existing interfaces. They were free to perform any gestures they believed would result in the activity visualised. They had complete freedom of movement, and were not given limitations in terms of using one or both hands, fingers, arms etc. However they had to “mime” the activity on the screen, in a sense that they had to perform the gesture at the same time the activity was shown on the screen for the third time. Each animation was three seconds long, in order to get an instant and intuitive reaction from the participants. They were also not told what the purpose of the study was until after all the activities were complete, to reduce bias. Experiments in the related literature are usually not explicitly time constrained. Cash and Maier (2016) observe designers working in a team, an uninterrupted design process, without introducing prompts or time limitations. However, designing in group would have set a pace naturally. Hurtienne et al. (2010), did not specify if individual gestures which tested use of primary metaphors within free-form gestures were time constrained. The participants were seated at one end of the table, and their activities were recorded using a LEAP Motion sensor (LEAP MOTION INC., 2017) secured to the table under their arms, as shown in Figure 1, and a camera positioned at the opposite end of the table, under the screen.



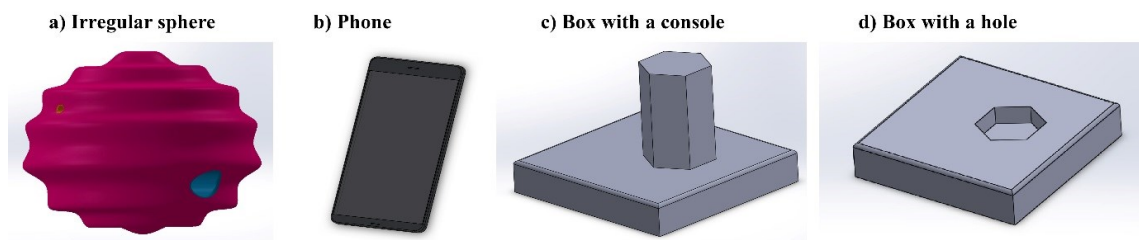
**Figure 1. Participant taking part in the study**

It was found that the use of the LEAP Motion sensor was not a reliable recording method, as the range of recording it supported was too small for the requirements of the study. Only 7 out of 69 gesture recordings were complete, and the rest were either partially recorded (e.g. only the bottom part of the hand was detected) or completely undetected by the device. In order to discover the most intuitive

gestures, the participants' attention was directed to the screen the videos of activities are being displayed on, which lead them to perform the gestures partially or fully outside of the LEAP sensor range. This could be remedied by allowing the participants to see if the sensor is detecting their gestures or not, in one corner of the screen, but that could on one hand, distract them from the activity shown on the screen, and on the other, potentially limit their gestures, as they could feel restricted to the LEAP sensor range. Therefore, we did not include the findings from the LEAP motion sensor recordings in the remainder of this paper, as LEAP did not provide sufficient quality or range of recording. The findings presented are those collected using the video camera recordings. One camera was used in the study and this was found to be sufficient to identify the gestures performed, except in two cases, where a second camera to identify the depth of movement would have been beneficial. If this study were to be replicated, a set up using two cameras would be preferable, first camera positioned in front of the participant to be used for the majority of gesture identification and second placed at 90 degrees to the side to identify depth of motion when needed. The remaining 67 gestures were easily identified and coded using these recordings. Videos of 3D object modification or manipulation were shown on a large screen. More sophisticated VR systems were considered initially, but use of 2D screen meant that the study could be performed at most standard workplaces, which have a screen a video can be projected to using a laptop.

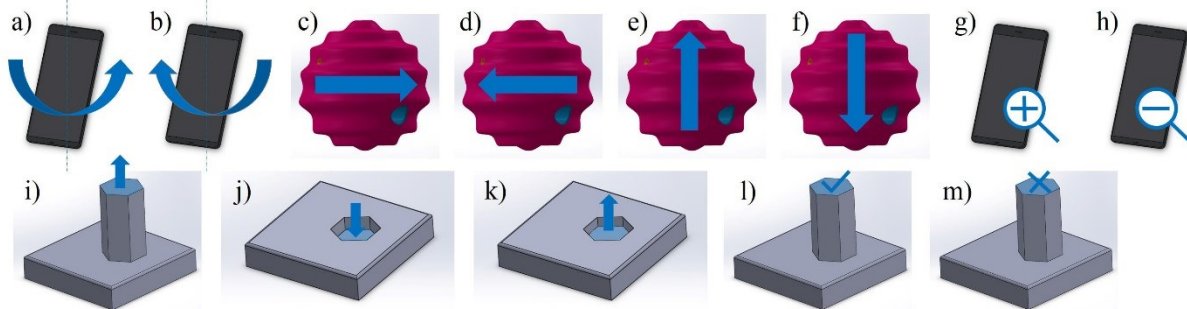
Seven participants were asked to take part in the study, possessing a minimum of five and a maximum of ten years of product design experience. All have completed a product design course, or a very closely related undergraduate course including a number of design classes, and are currently design focused PhD students or post-doctoral researchers. Their ages range from 24 to 29. Four participants were male and three were female. All were right handed.

Figure 2 shows three 3D objects: irregular sphere (a), phone (b) and box with a console (c)/hole (d).



**Figure 2. 3D models which were manipulated**

The objects chosen were geometrically simple, and did not include high level of detail, in order to minimise the time required for participants to perceive their shape upon viewing. The irregular sphere and the box do not have a specific recognisable use, while the model of a mobile phone does. These choices were intentional as the goal was to observe if familiarity of the participants with the object influences the interaction with it in any way, that is, if the way a phone would be handled in real life would be reflected in the way participants interact with a virtual representation of a phone. Types of activities that can be performed using these objects were also a factor in choosing the specific object for inclusion. For example, the phone and the sphere did not have surfaces that could be easily extruded or cut from the objects that were clearly visible, so instead the box with the console/hole was used for that. The objects were then animated in 3D representation to visualise the 13 activities shown in Figure 3.



**Figure 3. Activities performed on 3D objects; a) Rotation (clockwise), b) Rotation (counter clockwise), c) Panning (to the right), d) Panning (to the left), e) Panning (up), f) Panning (down), g) Zoom in (+), h) Zoom out (-), i) Extrusion of the console from the box, j) Insertion of the console into the hole of the box, k) Extrusion of the console from the box, l) Insertion of the console into the hole of the box, m) Extrusion of the console from the box.**

**Panning (down), g) Zooming (in), h) Zooming (out), i) Extrude, j) Extrude cut, k) Extrude cut shallower, l) Select, m) Deselect**

The sphere and the phone models were used to test manipulative gestures, activities shown in Figure 3a-h, and a familiar shape (the phone) and a shape without an implied use (the sphere) were used in order to test if the gestures used to manipulate the object would differ. The box with a console/hole was used for activities shown in Figure 3i-m, and the shape was a generic shape used to demonstrate extrusion of a part of the object, removal of a part of an object, selection and deselection. Extrude cut activity had two variations, based on the direction of the extrusion. Extrude cut shallower was reducing the volume being extruded, whereas extrusion cut was increasing it. Each participant was asked to perform gestures to complete eleven activities: two rotations (one with each model type), two panning activities (one left or right, and one up or down, one with each model type), two zoom activities (one with each model type), one extrusion, one extrude cut activity, one selection and one deselection. This choice was made to minimise the number of activities per participant, so that the length of duration of each segment of the study could be kept at the minimum and ensure the participants' full focused attention, while at the same time collecting sufficient amount of data to draw conclusions from.

#### **4. Findings**

Camera recordings were reviewed and coded by two coders. First, gestures were classified based on the paths traced by participants' hands and grouped by similarity. For example for zooming in four groups of gestures were identified: pulling the object up, pulling the object back, pulling two fingers apart to stretch the object and pushing the object to the front. Then these were coded to note the use of hands. Then the gestures were coded to note if a grasping motion was noticed in any part of the gesture. Finally, gestures were coded to ascertain if the participants were interacting with virtual objects as they would with a physical 3D object suspended in the air in front of them. That is, if hand motions performed mimicked the three-dimensional motions of the object displayed on screen, or if they were performing gestures which indicated interaction with a 2D screen. Interaction with a 2D screen would be suspected if gestures were performed in a plane even if objects on the screen were moving three-dimensionally. While statistical analysis can show statistically significant results for small samples, they are subject to high margins of uncertainty (Button et al., 2013, Hay et al., 2017). This is why sample sizes for the first two hypotheses, where there are between samples comparisons, were deemed too small to justify statistical analysis, and the findings for them are descriptive. Hypotheses three and four were statistically analysed, as due to the fact only one factor was observed for them the samples were comparatively larger. For hypotheses three and four inter-coder reliability was calculated using Krippendorff's Alpha reliability estimate. For hypothesis three the coders agreed in 100% of the cases, and alpha value was 1. For hypothesis four they agreed in 83% of the cases, and the alpha value was 0.65. This alpha value is not high enough to assume sufficient level of agreement, and cases where the coders disagreed were examined in order to identify the cause of disagreement. This is detailed in section 4.4.

##### **4.1. Findings on gestures per activity**

Focusing on the type of gestures performed for the same activities, results analysis has shown that while a number of different gestures appear for the same activities, majority of them are repeated by more than one participant.

Table 1 shows that majority of participants have, without being prompted to do so, performed one of the two gestures that have most frequently arisen for each of the activities. They did not always use the same combination of hands. Use of hands is indicated by the column number of gestures was placed in, where L indicates use of left hand, R use of right hand, B use of both hands (where both hands were used performing the same or mirrored movement), RD use of both hands but right hand used for the dominant movement, and LD use of both hands used but left hand used for the dominant movement.



Some participants performed gestures that no other participants repeated, and quantities of those are shown in the “number of gestures that were not repeated” column.

**Table 1. Gestures performed by different participants for the same activities**

Gesture	Different participants performed the same gesture for the same activity										Number of gestures that were not repeated	Total
	1 <sup>st</sup> type of gesture					2nd type of gesture						
	L	R	B	BLD	BRD	L	R	B	BLD	BRD		
Pan down		2	3								1	6
Pan up											2	2
Pan right	1	2										3
Pan left			2									3
Rotation cw	1	1	2								3	7
Rotation ccw		2	1						2	1	1	7
Zoom in		2	2								3	7
Zoom out		4	2								1	7
Select surface				2	2 <sup>T</sup> +1 <sup>F</sup>		2					7
Deselect surface				1 <sup>T</sup> +1 <sup>F</sup>	1 <sup>T</sup> +2 <sup>F</sup>		2					7
Extrude cut				1	2							3
Extrude cut shallower					3						1	4
Extrude up				2	3						1	6

For the selection and deselection of a surface, some participants used slightly different motions. F in the superscript denotes that the hand is performing a flicking motion, T that it is performing a tapping motion. Occasionally one finger or three fingers were used instead of the entire hand, but as the gestures were of the same nature, differentiation between the numbers of fingers used was not made at this stage. Although no instructions were given to the participants prior to the study, repetition in gestures between them was noticeable for most of the activities. Nine activities were most frequently performed using one type of a gesture e.g. to pan down five participants have used the gesture shown in Figure 4a. For three of the activities two different types of gestures are repeated e.g. to select a surface four participants perform the gesture shown in Figure 4b, and two participants perform the gesture shown in Figure 4c.



**Figure 4. Examples of repeated gestures for pan down activity (a), and select activity (b and c)**

The exception were gestures used for the pan up activity, where two different gestures were performed by two different participants. Due to asking the participants to perform only two panning activities in total and random assignment of the four variants of the panning activity, their totals differ, and larger number of repetitions per specific panning gesture would be required to reach reliable conclusions. Similarly, the totals on the extrude type gestures vary, as one participant was only asked to perform one extrude cut gesture, so they add up to 7. One participant failed to complete the extrude up gesture hence only six are reported.

#### 4.2. Findings on repeated gestures per participant

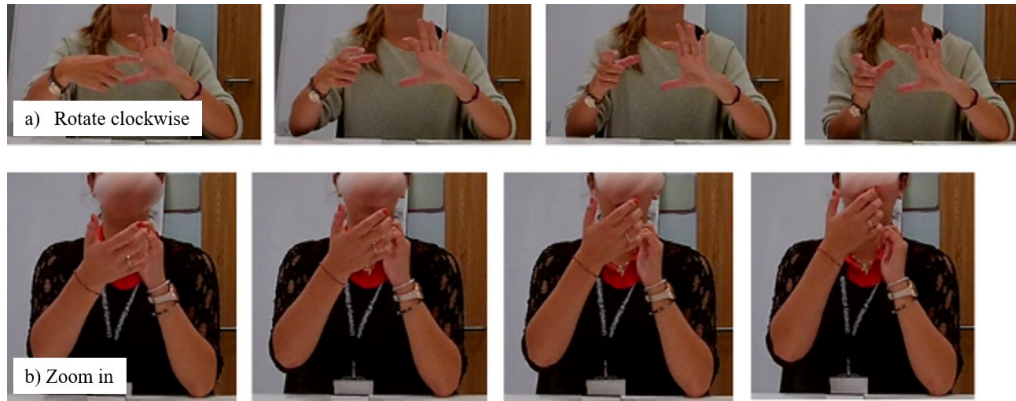
Observations on the repetition of gestures for the same type of activity by the same participants for different 3D objects shows that majority of participants do use same gestures for the same activities. For zooming, panning, and rotation activity 71% of participants use the same type of a gesture (while they employ different hands or different number of fingers to manipulate objects in 40% of those for panning activity, and 60% for zooming and rotation activity). For extrude cut and extrude up activity 100% of participants performed the same type of a gesture. For the selection and deselection activity, 57% participants used same arrangement of hands for both types of activities, and 43% used a partial version of the gesture (dominant hand was used for selection, but object was not held with the other hand). When both hands are used 80% of participants used the flicking gesture to select/deselect, and when only one hand was used 100% of participants used the tapping gesture. Table 2 provides an overview of the number of repeated gestures per participant.

**Table 2. Number of repeated gestures for the same type of activity by the same participant**

Gesture	Same gesture same hand use	Same gesture different combination of hands used	Different gestures	Total
Panning	2	3	2	7
Rotation	3	2	2	7
Zooming	2	3	2	7
Extrusion	6			6
Select/deselect	1 <sup>T</sup> +2 <sup>F</sup>	1 <sup>F</sup>	3 <sup>T</sup>	7

#### 4.3. Interaction with 3D objects shown on a 2D screen

Videos of 3D object manipulation and modification were shown on a 2D screen. Considering how ubiquitous 2D screens are in daily use, e.g. phones, tablets, touchscreens, some legacy bias was anticipated (Morris et al., 2014). Interactions that participants were familiar with, from previous experience, might have been used while viewing 3D objects displayed on a 2D screen. However, over 80% of the gestures each participant performed showed interaction with the virtual object was arguably performed as it would be with a physical object. When all gestures performed are put together, 94% of them were performed out of the bounds of a vertical 2D plane the objects were shown in. For example, majority of rotations were performed by participants holding the axis of rotation of an object in place with one hand, and using the other hand to rotate the object around it, moving in a circular motion in a plane perpendicular to the imagined axis, as can be seen in Figure 5a. Zooming motions were another example. In interaction with 2D surfaces typically tips of two fingers are moved closer together, or further apart, following a straight line in a plane to indicate zooming in or zooming out. Majority of participants instead pulled or pushed the imaginary 3D object to zoom in or out, as shown in Figure 5b.



**Figure 5. Examples of three-dimensional interaction**

Chi-square goodness of fit was used to test if this could have happened by chance. The null hypothesis for this test was that roughly the same number of gestures would be interacted with as if they were perceived as 2D objects on a screen or 3D objects suspended in air in front of the participants ( $H_0$ ). Table 3 displays observed and expected number of counts for 2D based and 3D based interaction with virtual 3D objects.

**Table 3. Observed counts of 2D and 3D interactions with a virtual model**

Observed counts	Count of 2D/3D
2D	3
3D	66
<b>Grand total</b>	<b>69</b>

$$X^2 = \sum \frac{(O-E)^2}{E} = 57.522; \quad (1)$$

Where O – observed frequency, E – expected frequency.

$$v = (\text{number of categories after pooling}) - (\text{number of parameters estimated}) - 1 = 1$$

Expected frequency of 2D and 3D interaction was 34.5, while encountered frequency was 3 for 2D interaction and 66 for 3D interaction. The obtained Chi-Square value (57.522) has a p of 0.000: this is  $< 0.05$ , and the conclusion is that the observed frequencies of use of 3D based interaction are significantly different from what would be expected to happen by chance. This indicates that effects of legacy bias are insignificant.



#### 4.4. Effect of 3D object recognisability on the nature of interaction

The 3D objects shown in the videos were a spherical shape, a mobile phone, and a square box, which had a hexagonal profile either extruded from it, or cut from it. The mobile phone was assumed to be the only object which participants would have the experience of handling habitually. The goal of the study was to test if familiarity with the objects would lead the participants to interact with different shapes in different ways. It was assumed that use of a grasping gesture would indicate that shape is recognisable, as to turn, push away or otherwise handle a mobile phone in reality, users would first need to grasp it to secure a hold of it. The null hypothesis of the study was that there would be no significant association between the shape of the object and grasping i.e. type of the interaction with it. Chi-square test of association between two independent variables was used, and number of gestures including grasping are shown in Table 4.

**Table 4. Summary of number of gestures which included or did not include grasping**

Observer	Grasp	Sphere	Phone	Box	Total
1	Y	12	13	13	39
1	N	9	14	8	30
2	Y	10	19	13	42
2	N	11	8	8	27

A 2x2 table had  $(3-1) \times (2-1) = 2$  degrees of freedom.

$$X_{\text{coder1}}^2 = \sum \frac{(O-E)^2}{E} = 0.956; \quad (2)$$

$$X_{\text{coder2}}^2 = \sum \frac{(O-E)^2}{E} = 2.581; \quad (3)$$

For coder one Chi-Square is 0.956, with two degrees of freedom, has a p of 0.620. For coder two Chi-Square is 2.581, with two degrees of freedom, has a p of 0.257. Both mean the null hypothesis cannot be rejected, and there is no proof of significant association between shape and grasping.

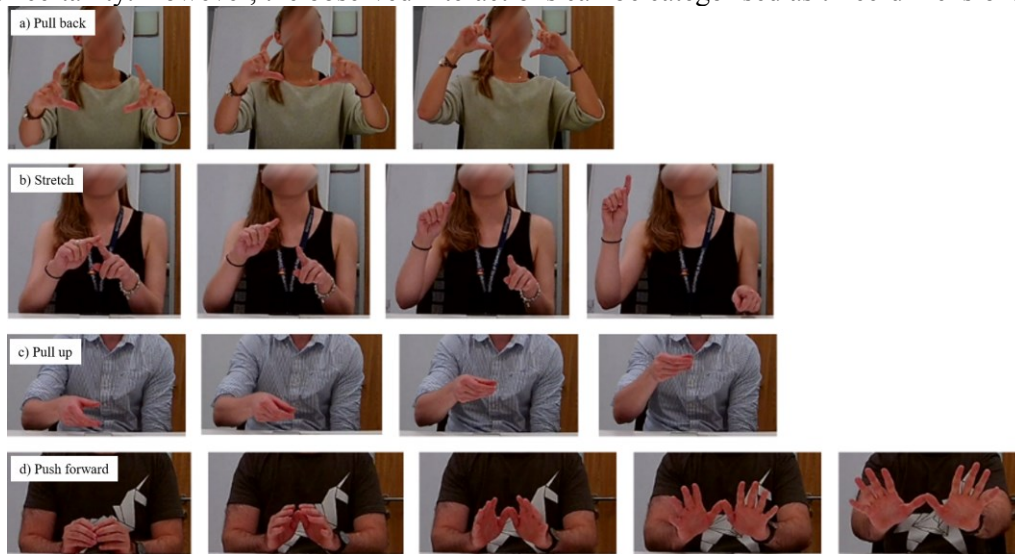
Grasping is not the only action which could be used to characterise a type of interaction with a recognisable object, and if a larger scale study was to be performed different characterisation features should be sought.

Where coders disagreed on existence of grasp, the cases were reviewed and discussed. First coder required the shape of the hand to be in a form of a grasp, as if the user was picking up a physical object. Second coder assumed a grasp even if this was not present, but it was clear that a user was holding a part of the object down with an open hand. For example, base of the parts i-m was often held down with one open palm, while the other hand was used to manipulate the console or the cut. Second source of disagreement were tasks performed with one hand where grasp was not significant enough to assume that if a physical object was handled the user would be able to lift it. In these cases, first coder assumed grasp was present, and the second did not. This exemplifies the need for definition of additional measures of interaction with an object which would indicate object recognisability.

## 5. Discussion

There were concerns about the perception of 3D shapes shown on the 2D screen, particularly depth perception and if the interaction with the object would be more akin to that encountered in tablet or phone interaction. For example, to zoom in on a tablet or a phone a user usually slides tips of two fingers apart in a straight line, or pinches the screen. It was found that this does not seem to be an issue. The participants did not interact with it in a vertical plane the object was shown in, but instead used three-dimensional space in front of them. The majority of participants interacted with the virtual objects as if they were physical objects suspended in the air in front of them. Using the zoom example, majority of the participants “pulled” the object back to zoom (Figure 6a), while one participant used the phone

screen based zoom activity (Figure 6b), one pulled the object up (Figure 6c), and one pushed it to the front (Figure 6d). This is encouraging, as it indicates that participants seem to have a good perception of three-dimensionality of an object even if they see it on a two-dimensional screen. The participants were not asked if they perceived objects as 3D objects, so it is impossible to claim what their perception was with certainty. However, the observed interactions can be categorised as three-dimensional.



**Figure 6. Examples of the zoom in gesture**

While it cannot be concluded that 2D representation of the objects did not have any effect on the gestures participants have performed, the gestures performed tended to use the three-dimensional space rather than single plane the object was shown in, and this likely means that for initial gesture elicitation inexpensive 2D screens and cameras could be used for gesture identification.

Although this study included only seven participants, some conclusions can be drawn on instances of repetition of same gestures used to perform same actions. The participants were not given any guidance in terms of how to perform gestures. They were asked to use their fingers, hands and/or arms to achieve the activity they have previously seen happen on the screen twice in a row. It was noticed that there was repetition of gestures. Depending on the specific gesture 50-100% of gestures a participant was performing for the same activity for different shapes were repeated. Observing the gestures different participants performed for the same activity, aside from two activities (counter clockwise rotation where two different types of gestures were repeated, and zoom in where there were three different gestures performed in addition to the repeated gesture), more than 70% of the same gestures were performed by different participants for the same activity. While this is not proof that standard gestures for 3D object manipulation and modification exist, it does support the argument for a necessity of further exploration of natural and intuitive gestures for conceptual design.

The nature and appropriateness of the gestures for the activity they were used for was not explored, as the sample sizes were small and even if the findings could have statistical significance, they would be subject to high margins of uncertainty. However, if the sample size was significantly increased further research into appropriateness of the gestures for the activity could be performed.

In addition to the small scale of the study, there are additional limitations, which may have affected the results. One of the participants had failed to perform one of the extrude cut gestures. As gestures are observed as individual events, this was not an issue during the data analysis, but it did mean that for extrude cut, and extrude cut shallower activities there was one less pair of gestures to observe when reporting on repeated gestures per participant. Grasping was used as an indicator of specific type of an interaction with an object, which had a recognisable purpose, or was handled more frequently by users in reality. While grasping is one of the characteristics of a unique type of interaction, use of grasping could also be affected by the shape itself, independently of its use. It could also reflect a personal preference of the participant. Additionally, grasping may not be the only indicator of the recognisability of an object, as differences between different coders perceptions of object interaction would suggest,

and interaction with differently shaped objects with different purposes should be explored. This could then lead to identification and definition of additional indicators of interaction with a recognizable object. The study was coded by two researchers. While parameters for coding were defined prior to the study being conducted, for the aspects of coding which allowed for more subjective interpretation and more variance within the parameters there were disagreements. In the future this could be mitigated by increasing the number of coders, and providing more detailed coding parameters.

## 6. Conclusion

This study has demonstrated that designers observing a 3D object on a 2D screen interact with the object as if it was a physical object suspended in air in front of them. While it cannot be claimed this indicated they perceived it as a 3D object, it was noticeable that there was almost no legacy bias stemming from the use of phones or tablets. This would mean that if 3D representation technology such as VR is not available, 2D screens could be an acceptable replacement for the related studies. Designers have performed similar gestures for the same activities performed on different objects both between and within subjects. While the sample of the study is not large enough to conclusively confirm or deny the effect of this on design, it suggests further research would be beneficial in order to identify which gestures are the most natural and most intuitive for manipulation and modification of 3D objects during conceptual design. The difference in object shapes and their recognisability however, has not been demonstrated to affect the nature of a designer's interaction with it. A larger study would be required to achieve conclusive findings on the nature of intuitive and unguided gesture interaction with virtual 3D objects, but the small-scale study findings given in this paper show promise.

## Acknowledgment

This research was made possible by funding from University of Strathclyde and insights from design research group at Department of Design Manufacturing Engineering Management.

## References

- Anonymous (2017). Details omitted for double-blind reviewing.
- Arroyave-Tobón, S., Osorio-Gómez, G. & Cardona-McCormick, J. F. (2015), "AIR-MODELLING: A tool for gesture-based solid modelling in context during early design stages in AR environments", *Computers in Industry*, Vol. 66, pp. 73-81. doi: 10.1016/j.compind.2014.10.007
- Asadzadeh, P., Kulik, L. & Tanin, E. (2012). "Gesture recognition using RFID technology", *Personal and Ubiquitous Computing*, Vol. 16 No. 3, pp 225-234. doi: 10.1007/s00779-011-0395-z
- Beattie, N., Horan, B. & McKenzie, S. (2015), "Taking the LEAP with the Oculus HMD and CAD - Plucking at thin Air?", *Procedia Technology*, Vol. 20, pp. 149-154. doi: 10.1016/j.protcy.2015.07.025
- Bourdot, P., Convard, T., Picon, F., Ammi, M., Touraine, D. & Vézien, J. M. (2010), "VR-CAD integration: Multimodal immersive interaction and advanced haptic paradigms for implicit edition of CAD models", *Computer-Aided Design*, Vol. 42 No. 5, pp. 445-461. doi: 10.1016/j.cad.2008.10.014
- Buchmann, V., Violich, S., Billingham, M. & Cockburn, A. (2004), "FingARtips: gesture based direct manipulation in Augmented Reality", *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, 2004, ACM, pp. 212-221. doi: 10.1145/988834.988871
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. & Munafò, M. R. (2013), "Power failure: why small sample size undermines the reliability of neuroscience", *Nature Reviews Neuroscience*, Vol. 14, pp. 365-376. doi: 10.1038/nrn3475
- Cash, P. & Maier, A. (2016), "Prototyping with your hands: the many roles of gesture in the communication of design concepts", *Journal of Engineering Design*, Vol. 27 no. 1-3, pp. 118-145. doi: 10.1080/09544828.2015.1126702
- Dani, T. H. & Gadh, R. (1997), "Creation of concept shape designs via a virtual reality interface", *Computer-Aided Design*, Vol. 29 No. 9, pp. 555-563. doi: 10.1016/S0010-4485(96)00091-7
- Gao, Z. & Gibson, I. (2006), "Haptic sculpting of multi-resolution B-spline surfaces with shaped tools", *Computer-aided design*, Vol. 38 No. 6, pp. 661-676. doi: 10.1016/j.cad.2006.02.004
- Han, Y. C. & Han, B.-J. (2014), "Virtual pottery: a virtual 3D audiovisual interface using natural hand motions", *Multimedia Tools and Applications*, Vol. 73 No. 2, pp. 917-933. doi: 10.1007/s11042-013-1382-3

- Hay, L., McTeague, C., Duffy, A. H., Pidgeon, L. M., Vuletic, T. & Greal, M. (2017), "A systematic review of protocol studies on conceptual design cognition", *Design Computing and Cognition'16*. Springer. doi: 10.1007/978-3-319-44989-0\_8
- Hürst, W. & Van Wezel, C. (2013), "Gesture-based interaction via finger tracking for mobile augmented reality", *Multimedia Tools and Applications*, Vol. 62 No. 1, pp. 233-258. doi: 10.1007/s11042-011-0983-y
- Hurtienne, J., Stöbel, C., Sturm, C., Maus, A., Rötting, M., Langdon, P. & Clarkson, J. (2010), "Physical gestures for abstract concepts: Inclusive design with primary metaphors", *Interacting with Computers*, Vol. 22 No. 6, pp. 475-484. doi: 10.1016/j.intcom.2010.08.009
- Leap Motion Inc. (2017), *Leap Motion Inc.* [Online]. Available at: <https://www.leapmotion.com/> (Accessed 27.11.2017).
- Kim, Y. & Park, J. (2014), "Study on interaction-induced symptoms with respect to virtual grasping and manipulation", *International Journal of Human-Computer Studies*, Vol. 72 No.2, pp. 141-153. doi: 10.1016/j.ijhcs.2013.09.011
- McNeill, D. (1985), "So you think gestures are nonverbal? *Psychological review*, Vol. 92 No. 3, pp. 350-371. doi: 10.1037/0033-295X.92.3.350
- Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., Schraefel, M. C. & Wobbrock, J. O. (2014), "Reducing legacy bias in gesture elicitation studies", *interactions*, Vol. 21 No. 3, pp. 40-45. doi: 10.1145/2591689
- Morris, M. R., Wobbrock, J. O. & Wilson, A. D. (2010), "Understanding users' preferences for surface gestures", *Proceedings of Graphics Interface 2010*, Canadian Information Processing Society, pp. 261-268.
- Noor, A. K. & Aras, R. 2015. "Potential of multimodal and multiuser interaction with virtual holography", *Advances in Engineering Software*, Vol. 81, pp. 1-6. doi: 10.1016/j.advengsoft.2014.10.004
- Qin, S., Wright, D. K., Kang, J. & Prieto, P. A. (2006), "Use of three-dimensional body motion to free-form surface design", *Proceedings of the Institution of Mechanical Engineers*, Vol. 220 No. 2, pp. 335-339. doi: 10.1243/095440506X77616
- Quek, F. (2004), "The catchment feature model: A device for multimodal fusion and a bridge between signal and sense", *EURASIP Journal on Applied Signal Processing*, Vol. 2004, pp. 1619-1636. doi: 10.1155/S1110865704405101
- Schmidt, A. (2015), "Following or leading?: the HCI community and new interaction technologies", *interactions*, Vol. 22 No. 1, pp. 74-77. doi: 10.1145/2692980
- Shankar, S. S. & Rai, R. (2014), "Human factors study on the usage of BCI headset for 3D CAD modeling", *Computer-Aided Design*, Vol. 54, pp. 51-55. doi: 10.1016/j.cad.2014.01.006
- Song, J., Cho, S., Baek, S.-Y., Lee, K. & Bang, H. (2014), "GaFinC: Gaze and Finger Control interface for 3D model manipulation in CAD application", *Computer-Aided Design*, Vol. 46, pp. 239-245. doi: 10.1016/j.cad.2013.08.039
- Vinayak & Ramani, K. (2015), "A gesture-free geometric approach for mid-air expression of design intent in 3D virtual pottery", *Computer-Aided Design*, Vol. 69, pp. 11-24. doi: 10.1016/j.cad.2015.06.006
- Wobbrock, J. O., Morris, M. R. & Wilson, A. D. (2009), "User-defined gestures for surface computing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2009*. ACM, pp. 1083-1092. doi: 10.1145/1518701.1518866
- Wright, M., Lin, C.-J., O'Neill, E., Cosker, D. & Johnson, P. (2011), "3D gesture recognition: an evaluation of user and system performance", *International Conference on Pervasive Computing, 2011*, Springer, pp. 294-313. doi: 10.1007/978-3-642-21726-5\_19
- Xu, Z., Li, Y. & Su, Z. (2009), "Research on computer aided creative design", *Computer-Aided Industrial Design & Conceptual Design, CAID & CD 2009. IEEE 10th International Conference on*, IEEE, pp. 1203-1208. doi: 10.1109/CAIDCD.2009.5375209
- Yeo, H.-S., Lee, B.-G. & Lim, H. (2015), "Hand tracking and gesture recognition system for human-computer interaction using low-cost hardware", *Multimedia Tools and Applications*, Vol. 74 No. 8, pp. 2687-2715. doi: 10.1007/s11042-013-1501-1
- Zhu, W. (2008), "A methodology for building up an infrastructure of haptically enhanced computer-aided design systems", *Journal of Computing and Information Science in Engineering*, Vol. 8 No. 4. doi: 10.1115/1.2988340
- Zimmerman, T. G., Lanier, J., Blanchard, C., Bryson, S. & Harvill, Y. (1987), "A hand gesture interface device", *Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface, ACM SIGCHI Bulletin, 1987*. ACM, pp. 189-192. doi: 10.1145/1165387.275628